

**PATENT**Atty Docket No.: 10017990-1  
App. Scr. No.: 10/020,256**IN THE CLAIMS:**

Please find below a listing of all of the pending claims. The statuses of the claims are set forth in parentheses.

**1. (Currently Amended)** A video processing device, comprising:

an audio event detecting means for detecting audio events in a video data; and  
a memory communicating with said audio event detecting means and storing video data and audio data corresponding to said video data, said memory further storing a predetermined energy threshold, a predetermined ZCR variance threshold, and a predetermined ZCR amplitude span threshold;

wherein said audio event detecting means includes an energy detector, which detects an audio event in said audio data by measuring an energy content of said audio data and indexes said video data at about a beginning of said audio event and, wherein said audio event detecting means further includes a ZCR detector for generating a ZCR value from said audio data and comparing a ZCR variance and a ZCR value span of said audio data to the predetermined ZCR variance threshold and to said predetermined ZCR amplitude span threshold, respectively, if said energy content satisfies said predetermined energy threshold.

**2. (Previously Presented)** The device of claim 1, wherein said audio event detecting means indexes said video data by extracting and storing one or more representative video frames.**3. (Previously Presented)** The device of claim 1, wherein said audio event detecting means indexes said video data by inserting index data into said video data.**4. (Previously Presented)** The device of claim 1, wherein said audio event detecting means indexes said video data by saving one or more index pointers.**5. (Previously Presented)** The device of claim 1, wherein said audio event detecting means indexes said video data by recording one or more time stamps.**6. (Original)** The device of claim 1, wherein said audio event comprises speech.

PATENT

Atty Docket No.: 10017990-I  
App. Ser. No.: 10/020,256

7. (Original) The device of claim 1, wherein said audio event comprises music.
8. (Original) The device of claim 1, wherein said video processing device comprises a video recorder device.
9. (Original) The device of claim 1, wherein said video processing device comprises a video editor device.
10. (Original) The device of claim 1, wherein said video processing device comprises a video authoring device.
11. (Currently Amended) A video processing device, comprising:
  - a processor;
  - an audio event detector communicating with said processor, wherein said audio event detector includes an energy detector communicating with said processor and a ZCR detector communicating with said processor, and
  - a memory communicating with said processor, said memory storing video data and audio data corresponding to said video data, said memory further storing a predetermined energy threshold, a predetermined ZCR amplitude span threshold, and a predetermined ZCR variance threshold;

wherein said audio event detector detects an audio event in said audio data by measuring an energy content of said audio data, generating a ZCR value from said audio data, and comparing a ZCR variance and a ZCR value span of said audio data to the predetermined ZCR variance threshold and to the predetermined ZCR amplitude span threshold, respectively, if said energy content satisfies said predetermined energy threshold, and wherein said processor indexes said video data at about a beginning of said audio event.
12. (Original) The device of claim 11, wherein said video processing device comprises a video recorder device.

PATENT

Atty Docket No.: 10017990-1  
App. Ser. No.: 10/020,256

13. (Original) The device of claim 11, wherein said video processing device comprises a video editor device.

14. (Original) The device of claim 11, wherein said video processing device comprises a video authoring device.

15. (Original) The device of claim 11, wherein said processor indexes said video data by extracting and storing one or more representative video frames.

16. (Original) The device of claim 11, wherein said processor indexes said video data by inserting index data into said video data.

17. (Original) The device of claim 11, wherein said processor Indexes said video data by saving one or more index pointers.

18. (Original) The device of claim 11, wherein said processor Indexes said video data by recording one or more time stamps.

19. (Currently Amended) The device of claim 11, wherein said memory stores a predetermined energy threshold, a predetermined ZCR variance threshold, a predetermined ZCR amplitude span threshold, and a predetermined set of speech harmonics thresholds, and wherein said audio event detector further comprises:

a ZCR-detector-communicating-with-said-processor-and-generating-a-ZCR-value-from-said-audio-data;

a spectrum detector communicating with said processor and generating a frequency spectrum from said audio data;

wherein said audio event detector compares harmonic frequency components in said frequency spectrum to said predetermined set of speech harmonics thresholds and detects a speech audio event if said harmonic frequency components fall within said predetermined set of speech harmonics thresholds if [[a]] said ZCR value span exceeds satisfies said predetermined ZCR amplitude span threshold, if a variance between said ZCR value and one

PATENT

Atty Docket No.: 10017990-1  
App. Scr. No.: 10/020,256

or more previous ZCR values satisfies is above the predetermined ZCR variance threshold, and if said energy content satisfies is greater than said predetermined energy threshold.

20. (Original) The device of claim 19, wherein said spectrum detector comprises an FFT processor.

21. (Currently Amended) The device of claim 11, wherein said memory stores a predetermined energy threshold and a predetermined frequency change threshold and wherein said audio event detector further comprises:

a spectrum detector communicating with said processor and generating a frequency spectrum from said audio data;

a peak detector communicating with said processor and said spectrum detector, said peak detector receiving said frequency spectrum, detecting frequency peaks in said frequency spectrum and generating a frequency peak output; and

wherein said audio event detector compares frequency peaks in two or more frequency peak outputs and detects a music audio event if said frequency peaks in said two or more frequency peak outputs are substantially stable and if said energy content is greater than said predetermined energy threshold.

22. (Original) The device of claim 21, wherein said spectrum detector comprises an FFT processor.

23. (Original) The device of claim 21, wherein said music detector compares frequency peaks in two or more consecutive frequency peak outputs.

24. (Currently Amended) A method of indexing a video data, comprising the steps of:

detecting an audio event in an audio data corresponding to said video data by

comparing an energy content of said audio data to a predetermined energy threshold and

comparing a ZCR variance and a ZCR value span of said audio data to a predetermined ZCR variance threshold and to a predetermined ZCR amplitude span

PATENT

Atty Docket No.: 10017990-1  
App. Ser. No.: 10/020,256

threshold, respectively, if said energy content satisfies said predetermined energy threshold;  
and

indexing one or more representative video frames of said video data at about a beginning of said audio event.

25. (Original) The method of claim 24, with the step of detecting said audio event further comprising detecting a speech audio event in said audio data.

26. (Currently Amended) The method of claim 24, with the step of detecting said audio event further comprising the steps of:

comparing a ZGR-variance-and-a-ZGR-value-span-of-said-audio-data-to-a predetermined-ZGR-variance-threshold-and-to-a-predetermined-ZGR-amplitude-span threshold, respectively, if said energy content is greater than said predetermined energy threshold;

comparing harmonic frequency components of said audio data to a predetermined set of speech harmonics thresholds if said ZCR variance and said ZCR value span satisfies exceed said predetermined ZCR variance threshold and said predetermined ZCR amplitude span threshold, respectively; and

detecting a speech audio event if said harmonic frequency components are within said predetermined speech harmonics range.

27. (Original) The method of claim 24, with the step of detecting said audio event further comprising detecting a music audio event in said audio data.

28. (Previously Presented) The method of claim 24, with the step of detecting said audio event further comprising the steps of:

comparing frequency peaks in two or more consecutive frequency spectra if said energy content is greater than said predetermined energy threshold; and

detecting a music audio event if said frequency peaks in said two or more consecutive frequency spectra are substantially stable.

**PATENT****Atty Docket No.: 10017990-1  
App. Ser. No.: 10/020,256**

29. (Original) The method of claim 24, with the step of indexing further comprising indexing said video data at about a beginning of a semantically meaningful video scene.

30. (Original) The method of claim 24, with the step of indexing further comprising extracting and storing said one or more representative video frames.

31. (Original) The method of claim 24, with the step of indexing further comprising inserting index data into said video data.

32. (Original) The method of claim 24, with the step of indexing further comprising saving one or more index pointers.

33. (Original) The method of claim 24, with the step of indexing further comprising storing one or more time stamps.